JOURNAL OF MULTIVARIATE ANALYSIS 33, 89–105 (1990)

Sequential Confidence Sets with Guaranteed Coverage Probability and Beta-Protection

ISSA FAKHRE-ZAKERI

University of Maryland

Communicated by the Editors

A general procedure for constructing sequential confidence sets for a vector valued parameter, having a coverage probability at least $1-\alpha$ and probability of covering a certain set of false values at most β, is developed. The limiting values of the error probabilities are found as the parameter approaches the boundary points. Applications are made to the problem of confidence sets for the mean vector and the covariance matrix of a multivariate normal, and to the multiple regression model. © 1990 Academic Press, Inc.

1. INTRODUCTION

Let $\{P_\xi : \xi \in \Xi\}$ be a family of distributions. Suppose that it is desired to construct a confidence set for some function of ξ, say $\gamma(\xi)$, with values in an open space $\Gamma \subseteq \mathbb{R}^k$, based on the random sequence $\{\xi_n : n \geqslant 1\}$. The sequence of estimators $\{\xi_n : n \geqslant 1\}$, where ξ_n is based on n observations, will be assumed consistent for ξ whenever the distribution of $\{\xi_n : n \geqslant 1\}$ is generated by P_ξ. The vector valued function $\gamma = \gamma(\xi)$ is the parameter of interest and the rest of ξ may be regarded as a nuisance parameter.

Let $0 < \alpha < 1$ and let $\gamma_n = \gamma(\xi_n)$. A conficence set $S(\gamma_1, \gamma_2, ...,)$ based on the sequence $\{\xi_n : n \geqslant 1\}$ is said to have coverage probability at least $1-\alpha$ for $\gamma \in \Gamma$ if

$$P_\xi[\gamma \in S(\gamma_1, \gamma_2, ...,)] \geqslant 1-\alpha \qquad \text{for all} \quad \gamma \in \Gamma. \tag{1.1}$$

Let $0 < \beta < 1$ and for every $\gamma \in \Gamma$ let $F(\gamma)$ be a closed subset of Γ. We will say that F is a β-protection region for the confidence set $S(\gamma_1, \gamma_2, ...,)$ if

$$P_\xi[F(\gamma) \cap S(\gamma_1, \gamma_2, ...,) \neq \varnothing] \leqslant \beta \qquad \text{for all} \quad \gamma \in \Gamma. \tag{1.2}$$

Generally speaking, if γ is the true parameter, $F(\gamma)$, which is chosen by the

Received May 17, 1988; revised May 10, 1989.
AMS 1980 subject classifications: primary 62L10, 62F25; secondary 62H99, 62J05.
Key words and phrases: confidence sets, stopping time, multiple regression.

0047-259X/90 $3.00

experimenter, is the set of false points in the parameter space that should be excluded from the confidence set.

Given the error probabilities $0<\alpha<1$, $0<\beta<1$, and a family $\{F(\gamma), \gamma \in \Gamma\}$ of closed subsets of Γ, we want to construct a confidence set for γ such that it satisfies both (1.1) and (1.2). In (1.2) it will be assumed that F is such that there is no fixed sample size confidence set for γ satisfying both (1.1) and (1.2). Thus, we will consider sequential procedures.

Requirement (1.2) as a measure of precision was originally adopted by Wijsman [13] to obtain an upper $(1-\alpha)$ confidence interval for a normal mean μ with known variance. We should mention here that it may not always be the case in practice to demand a fixed-width confidence interval. Such would be the case if, for instance, μ represents the average performance of a certain piece of equipment and one is only interested in a lower bound. In this connection it is, perhaps, worth pointing out that often the desired precision is not fixed and may depend on the true value of the unknown parameter γ, for instance, if γ represents the toxicity of a drug or the intensity of a radioactive substance.

Considering that the length cannot serve as a measure of precision, requirement (1.2) is an alternative to the usual fixed-width (or fixed accuracy) confidence sets. This way of controlling the precision is especially appropriate for one-sided confidence intervals (confidence sets) that have infinite length (are unbounded), and in situations where the desired precision depends on γ.

This method has recently been applied to a variety of problems, for the mean μ of a normal population with known variance σ (Wijsman [13], [14]), the μ/σ in the normal population (Wijsman [15]), the mean of an exponential distribution (Juhlin [6]), and the mean of a distribution in the presence of nuisance parameters (Kim [7]). These results are limited, however, to the one-sided confidence interval with a specific β-protection region.

In the present paper we provide a general method that is capable of generating a wide variety of sequential confidence sets and that is applicable to different choices of β-protection regions. This method is also applicable whether the problem involves a nuisance parameter or not and includes the indicated results as special cases. To attain that, we will introduce a Borel measurable function on $\Gamma \times \mathbb{R}^k$ and a q-dimensional interval, $q \geqslant 1$, where the choice of the function and the interval will determine the type of confidence set and, consequently, the stopping time and the terminal decision rule.

The paper is organized as follows. Section 2 presents the general formulation and the assumptions. Section 3 contains the main theorems. Finally, Section 4 gives the applications to the multivariate normal with unknown mean and covariance matrix and to the general linear model.

2. PRELIMINARIES, ASSUMPTIONS

It will be assumed throughout that the parameter space Ξ is of the form $\Xi = \Gamma \times \Theta$, where $\gamma \in \Gamma \subset \mathbb{R}^k$ is the parameter of interest and $\theta \in \Theta \subset \mathbb{R}^l$ is the nuisance parameter. This does not change anything essential and simplifies notation and statements of the results. Let H be a Borel measurable function on $\Gamma \times \mathbb{R}^k$ into $\mathbb{R}^q$, $q \geqslant 1$, and let D be a closed q-dimensional interval. Let $S(x) = \{\gamma \in \Gamma : H(\gamma, x) \in D\}$. The sequential procedure below is based on the fact that one constructs a stopping time N relative to $\mathscr{F}_n$, an increasing system of σ-algebras associated with the random processes $\xi_n = (\gamma_n, \theta_n)$ and then takes $S_N = S(\gamma_N)$ to be the terminal decision rule. A large class of terminal decision rules can be specified by this method. To illustrate this, some examples are given later in Section 3.

Let $\{F(\gamma) : \gamma \in \Gamma\}$ be a family of closed subsets of Γ such that if $y \in F(\gamma)$ then $H(y, \gamma) \notin D$. This condition implies that $F(\gamma)$ and $S(\gamma)$ are disjoint. Then F will be taken as our β-protection region. We assume that there exists a positive-definite real symmetric $k \times k$ matrix $\Sigma(\gamma, \theta)$, depending on (γ, θ), such that

$$\sqrt{n}\, \Sigma^{-1/2}(\gamma, \theta)(\gamma_n - \gamma) \xrightarrow{\mathscr{L}} N(0, I_k) \qquad \text{as} \quad n \to \infty,$$

where $\Sigma^{-1/2}(\gamma, \theta)$ is the positive square root of the inverse of $\Sigma(\gamma, \theta)$, I_k is the $k \times k$ identity matrix, and $\to^{\mathscr{L}}$ denotes convergence in distribution.

Generally speaking, to define a stopping time, one has to consider the following questions: how close $F(\gamma)$ and $S(\gamma)$ are, how close γ is to the boundary of $S(\gamma)$ and finally how "large" the covariance matrix is. Let $d(\gamma) = d(S(\gamma), F(\gamma))$, where $d(A, B)$ is the usual distance between sets A and B, and let $r(\gamma)$ be the radius of the largest ball centered at γ and contained in $S(\gamma)$. It should be mentioned that later assumptions on H and D will gaurantee that $d(\gamma)$ and $r(\gamma)$ are positive. Intuitively the stopping time will be large if the covariance matrix is "large" and, $r(\gamma)$ or $d(\gamma)$ are small. Let $\lambda_1(\gamma, \theta)$ be the largest eigenvalue of $\Sigma(\gamma, \theta)$ and let $h : \Gamma \times \Theta \to \mathbb{R}^+$ be such that

$$\frac{h(\gamma, \theta)}{\sqrt{\lambda_1(\gamma, \theta)}} \geqslant b, \qquad \forall (\gamma, \theta) \in \Gamma \times \Theta, \text{for some } b > 0.$$

Finally, let $g : \Gamma \times \Theta \to \mathbb{R}^+$ be such that $g(\gamma, \theta) \leqslant \operatorname{Min}(r(\gamma), d(\gamma))$ and $\tau(\gamma, \theta) = h(\gamma, \theta)/g(\gamma, \theta)$. Define the stopping rule $N_{c,(\gamma, \theta)} = N$ by

$$N = \text{least integer } n \geqslant r_0 + c^2 \tau^2(\gamma_n, \theta_n), \tag{2.1}$$

in which $r_0 \geqslant 0$ is the initial sample size and will stay fixed throughout, and $c > 0$ still has to chosen.

According to the requirements (1.1) and (1.2) the sequential procedure (N, S_N) has to satisfy the following conditions:

$$P_{\gamma,\theta}[\gamma \in S(\gamma_N)] \geqslant 1-\alpha, \qquad \forall(\gamma, \theta) \in \Gamma \times \Theta, \tag{2.2}$$

$$P_{\gamma,\theta}[S(\gamma_N) \cap F(\gamma) \neq \varnothing] \leqslant \beta, \qquad \forall(\gamma, \theta) \in \Gamma \times \Theta. \tag{2.3}$$

It should be pointed out that when $\xi = (\gamma, \theta)$ approaches some boundary points of the parameter space Ξ, $\tau(\gamma, \theta)$ and consequently the stopping time defined in (2.1) may tend to be very large. In order to deal with these problems a suitable compactification of the parameter space is required. We consider the compactification of the parameter space, which is denoted by $\bar{\Xi} = \overline{\Gamma \times \Theta}$, induced by the imbedding,

$$h: \mathbb{R}^{k+l} \to D^{k+l} \qquad \text{defined by} \quad x \mapsto x/(1 + \|x\|),$$

where $D^{k+l} = \{x \in \mathbb{R}^{k+l}: \|x\| \leqslant 1\}$ and $\|\ \|$ denotes the Euclidean norm.

For every $\xi \in \Xi$, let $\{Y_{n\xi}: n \geqslant 1\}$ be a sequence of random vectors in $\mathbb{R}^k$, and let $F_{n\xi}$ be the distribution function of $Y_{n\xi}$. Let Φ_k be the distribution function of $N(0, I_k)$.

DEFINITION 2.1. We say that $Y_{n\xi} \to^{\mathscr{L}} N(0, I_k)$ as $n \to \infty$ uniformly in ξ (unif ξ) if and only if

$$\operatorname*{Sup}_{\xi \in \Xi} \operatorname*{Sup}_{x} |F_{n\xi}(x) - \Phi_k(x)| \to 0 \qquad \text{as} \quad n \to \infty.$$

DEFINITION 2.2. A sequence $\{Y_{n\xi}: n \geqslant 1\}$ of real-valued random variables is said to be uniformly continuous in probability (ucip), unif ξ if and only if for every $\varepsilon > 0$ there exists a $\delta > 0$ for which

$$\operatorname*{Sup}_{\xi \in \Xi} P[\max_{0 \leqslant k \leqslant n\delta} |Y_{n+[k\delta],\xi} - Y_{n\xi}| \geqslant \varepsilon] < \varepsilon \qquad \text{for all} \quad n \geqslant 1.$$

A sequence $\{Y_{n\xi}: n \geqslant 1\}$ of random vectors is said to be ucip, unif ξ if and only if each component is ucip, unif ξ.

ASSUMPTION A. (i) *For every* $(\gamma, \theta) \in \Gamma \times \Theta$, $(\gamma_n, \theta_n) \to^{\text{a.e.}} (\gamma, \theta)$ *as* $n \to \infty$;

(ii) $\sqrt{n}\, \Sigma^{-1/2}(\gamma, \theta)(\gamma_n - \gamma) \to^{\mathscr{L}} N(0, I_k)$ *as* $n \to \infty$, unif (γ, θ);

(iii) *the sequence* $\{\sqrt{n}\, \Sigma^{-1/2}(\gamma, \theta)(\gamma_n - \gamma): n \geqslant 1\}$ *is* ucip, unif (γ, θ);

(iv) *the function* $H(\gamma, x)$ *on* $\Gamma \times \mathbb{R}^k$ *is continuous and* $H(\gamma, \gamma) \in D^0$, *where* D^0 *denotes the interior of* D;

(v) *for every* $\gamma \in \Gamma$, *there exists a closed ball* $V(\gamma) \subset \Gamma$ *centered at* γ *such that if* $x \in V(\gamma)$ *and* $y \in F(y)$ *then* $H(y, x) \notin D$, *the radius of* $V(\gamma)$ *can be written as* $\eta_\gamma d(\gamma)$ *and* $\inf_{\gamma \in \Gamma} \eta_\gamma = \eta_0 > 0$;

(vi) *let* ξ^* *stand for any boundary point of* $\bar{\Xi}$ *at which* $\tau(\xi_n) \to \infty$ *as* $\xi \to \xi^*$; *for all* $n = 1, 2, \ldots$, *there is at least one such* ξ^* *and we shall denote the set of such* ξ^**'s by* $\bar{\Xi}^*$.

(vii) *for every* $\xi^* \in \bar{\Xi}^*$, $\xi_N \to \xi^*$ *as* $\xi \to \xi^*$ *for any* $c > 0$;

(viii) $r(\gamma_N)/r(\gamma)$, $d(\gamma_N)/d(\gamma)$, $g(\xi_N)/g(\xi)$, *and* $h(\xi_N)/h(\xi) \to^{\text{a.e.}} 1$ *as* $c \to \infty$ unif.(γ, θ) *or as* $\xi \to \xi^*$ *for any* $c > 0$.

Assumption A(iv) *implies that the set* $S(\gamma)$ *is Borel measurable for all* $\gamma \in \Gamma$, *and also that* $\gamma \in S^0(\gamma)$ *and* $\gamma \in A^0(\gamma)$, *where*

$$A(\gamma) := \{x \in \mathbb{R}^k : H(\gamma, x) \in D\}. \tag{2.4}$$

Assumption A(v) *implies that* $F(\gamma) \cap S(x) = \emptyset$ *for all* $x \in V(\gamma)$. *Thus the set*

$$B(\gamma) := \{x \in \Gamma : S(x) \cap F(\gamma) = \emptyset\} \neq \emptyset. \tag{2.5}$$

One could replace A(vii) *and* A(viii) *by the following assumptions.*

A′(vii) $\xi_n \to \xi^*$ *as* $\xi \to \xi^*$, $n = 1, 2, \ldots$;

A′(viii) *for every* $\xi_0 \in \bar{\Xi}$, $r(\gamma_n)/r(\gamma)$, $d(\gamma_n)/d(\gamma)$, $g(\xi_n)/g(\xi)$, *and* $h(\xi_n)/h(\xi) \to 1$ *as* $n \to \infty$, $\xi \to \xi_0$.

Obviously, A′(vii) *implies* A(vii), *and, by the method of proof given in Lemma* 2.10 *(Fakhre-Zakeri* [5]*), one can see easily that* A′(viii) *implies* A(viii). *However, in certain examples (see Fakhre-Zakeri* [5]*), one cannot establish* A′(vii) *and* A′(viii), *but* A(vii) *and* A(viii) *remain true because of the behavior of the stopping time.*

3. Main Theorems

THEOREM 3.1. *Under Assumption* A, *for any given error probabilities* $0 < \alpha < 1$, $0 < \beta < 1$, *and integer* $r_0 \geqslant 0$, *there exists* $c_1 > 0$ *such that if* $c \geqslant c_1$, *then* $\alpha(c, \xi) \leqslant \alpha$ *and* $\beta(c, \xi) \leqslant \beta$ *for all* $\xi \in \Xi$.

For the next theorem, we assume that the following limits exist as $\xi \to \xi^* \in \bar{\Xi}^*$,

$$\begin{aligned}
b^*(\xi^*) &:= \lim_{\xi \to \xi^*} \frac{h(\xi)}{\sqrt{\lambda_1(\xi)}} \\
A^*(\xi^*) &:= \lim_{\xi \to \xi^*} (g(\xi))^{-1} (A(\gamma) - \gamma) \\
B^*(\xi^*) &:= \lim_{\xi \to \xi^*} (g(\xi))^{-1} (B(\gamma) - \gamma) \\
\Sigma^*(\xi^*) &:= \lim_{\xi \to \xi^*} (\lambda_1(\xi))^{-1} \Sigma(\xi),
\end{aligned} \tag{3.1}$$

where $A(\gamma)$ and $B(\gamma)$ are defined in (2.4) and (2.5), respectively, and $\xi=(\gamma, \theta)$.

THEOREM 3.2. *Under Assumption* A *and condition* (3.1), *for any given* $c>0$ *and integer* $r_0 \geqslant 0$, *we have*

(i) $\alpha(c, \xi) \rightarrow P[N(0, \Sigma^*(\xi^*)) \notin cb^*(\xi^*)\, A^*(\xi^*)]$

(ii) $\beta(c, \xi) \rightarrow P[N(0, \Sigma^*(\xi^*)) \notin cb^*(\xi^*)\, B^*(\xi^*)]$,

as $\xi \rightarrow \xi^*$, *where* $N(0, \Sigma^*(\xi^*))$ *stands for a normal random vector with mean zero and covariance* $\Sigma^*(\xi^*)$.

Before going to the proofs of these theorems we shall provide here some examples in order to illustrate the procedure.

EXAMPLE 1. Consider the following spherical confidence set centered at γ with β-protection region being outside of a ball, also centered at γ. Specifically, let r be a function on Γ into $\mathbb{R}^+$, let $H(\gamma, x)=\|\gamma-x\|-r(x)$, where $\|\ \|$ denotes the Euclidean norm, and let $D=(-\infty, 0]$. Let $F(\gamma)=\{x \in \mathbb{R}^k : \|\gamma-x\| \geqslant (1/\eta)\, r(\gamma)\}$, where $0<\eta<1$. Then $S(x)=\{\gamma \in \Gamma: \|\gamma-x\|-r(x) \leqslant 0\}$, $A(\gamma)=\{x: \|\gamma-x\|-r(x) \leqslant 0\}$, $B(\gamma)=\{x: \|\gamma-x\|+r(x)<(1/\eta)\, r(\gamma)\}$. Let $g(\xi)=(1-\eta)\, r(\gamma)$ and for definiteness consider $h(\xi)=\sqrt{\lambda_1(\xi)}$. Then,

$$
\begin{aligned}
(g(\xi))^{-1}(A(\gamma)-\gamma) &= (1-\eta)^{-1}(r(\gamma))^{-1}(A(\gamma)-\gamma) \\
&= (1-\eta)^{-1}\left\{\frac{x-\gamma}{r(\gamma)} : \|x-\gamma\|-r(x) \leqslant 0\right\} \\
&= (1-\eta)^{-1}\left\{z: \|z\|-\frac{r(\gamma+r(\gamma)\, z)}{r(\gamma)} \leqslant 0\right\}. \qquad (3.2)
\end{aligned}
$$

If $r(\gamma) \rightarrow 0$ as $\xi \rightarrow \xi^*$ and $r(x+y)/r(x) \rightarrow 1$ as $y \rightarrow 0$ for all x, then it follows from (3.2) that

$$(g(\xi))^{-1}(A(\gamma)-\gamma) \rightarrow (1-\eta)^{-1} B(0,1) \quad \text{as} \quad \xi \rightarrow \xi^*,$$

where $B(0,1)$ is just the Euclidean ball in $\mathbb{R}^k$ of radius one centered at the origin. Thus, by Theorem 3.2,

$$\alpha(c, \xi) \rightarrow P[N(0, \Sigma^*(\xi^*)) \notin B(0, c/(1-\eta))] \quad \text{as} \quad \xi \rightarrow \xi^* \text{ for any } c>0.$$

Similarly,

$$\begin{aligned}(g(\xi))^{-1}(B(\gamma)-\gamma) &= (1-\eta)^{-1}(r(\gamma))^{-1}(B(\gamma)-\gamma) \\ &= (1-\eta)^{-1}\left\{z: \|z\| + \frac{r(\gamma+r(\gamma)z)}{r(\gamma)} \leqslant \frac{1}{\eta}\right\} \\ &\to B\left(0, \frac{1}{\eta}\right) \quad \text{as} \quad \xi \to \xi^*.\end{aligned}$$

Thus,

$$\beta(c, \xi) \to P\left[N(0, \Sigma^*(\xi^*)) \notin B\left(0, \frac{c}{\eta}\right)\right] \quad \text{as} \quad \xi \to \xi^* \quad \text{for any} \quad c>0.$$

EXAMPLE 2. Let ψ be a mapping of Γ into $\mathbb{R}^k$ such that $\psi(\gamma)<\gamma$. (Here and throughout the vector inequalities are componentwise.) Let $F(\gamma)=\{x\in\mathbb{R}^k: x_i \leqslant \psi_i(\gamma) \text{ for some } i=1, ..., k\}$, and let ρ be a continuous function on Γ into $\mathbb{R}^k$ such that $\rho(\gamma)\in(\psi(\gamma), \gamma)$, a k-dimensional interval, for all $\gamma\in\Gamma$. Then, the function

$$H: \Gamma\times\mathbb{R}^k \to \mathbb{R}^k \qquad \text{defined by} \quad (\gamma, x)\mapsto \rho(x)-\gamma,$$

and the k-dimensional closed interval $D=\{x\in\mathbb{R}^k: x\leqslant 0\}$ will produce a one-sided confidence set. In this case,

$$\begin{aligned}S(\gamma) &= \{x\in\Gamma: \rho(\gamma)\leqslant x\}, \\ A(\gamma) &= \{x: \rho(x)<\gamma\}, \\ B(\gamma) &= \{x: \psi(\gamma)<\rho(x)\}.\end{aligned}$$

Let ψ_1 and ψ_2 be mappings of Γ into $\mathbb{R}^k$ such that $\psi_1(\gamma)<\gamma<\psi_2(\gamma)$, and let $F(\gamma)$ be the complement of the k-dimensional interval $(\psi_1(\gamma), \psi_2(\gamma))$. Also, let ρ_1 and ρ_2 be continuous functions on Γ into $\mathbb{R}^k$ such that $\rho_1(\gamma)\in(\psi_1(\gamma), \gamma)$ and $\rho_2(\gamma)\in(\gamma, \psi_2(\gamma))$ for all $\gamma\in\Gamma$. Then, the function

$$H: \Gamma\times\mathbb{R}^k \to \mathbb{R}^k\times\mathbb{R}^k \qquad \text{defined by} \quad (\gamma, x)\to(\rho_1(x)-\gamma, \gamma-\rho_2(x)),$$

and the $2k$-dimensional closed interval $(-\infty, 0]\times(-\infty, 0]\subset\mathbb{R}^k\times\mathbb{R}^k$ will produce the box-shaped confidence set. In this case

$$\begin{aligned}S(\gamma) &= \{x: \rho_1(\gamma)\leqslant x\leqslant\rho_2(\gamma)\}, \\ A(\gamma) &= \{x: \rho_1(\gamma)<x<\rho_2(\gamma)\}, \\ B(\gamma) &= \{x: \psi_1(\gamma)<\rho_1(x) \text{ and } \rho_2(x)<\psi_2(\gamma)\}.\end{aligned}$$

The proofs of Theorems 3.1 and 3.2 use the following lemmas.

LEMMA 3.1. *$N/c^2\tau^2(\xi) \to^{\text{a.e.}} 1$ as $c \to \infty$,* unif *$\xi = (\gamma, \theta)$ or as $\xi \to \xi^*$ for any $c > 0$.*

LEMMA 3.2. *$\sqrt{N}\,\Sigma^{-1/2}(\xi)(\gamma_N - \gamma) \to^{\mathscr{L}} N(0, I_k)$ as $c \to \infty$,* unif *ξ, or as $\xi \to \xi^*$ for any $c > 0$.*

LEMMA 3.3. *Let $\{A_n\}$ be a sequence of $m \times k$ matrices whose entries are bounded and let $\{X_n\}$ be a sequence of random vectors with values in $\mathbb{R}^k$ such that*

$$X_n \xrightarrow{\mathscr{L}} N(0, I_k) \qquad \text{as} \quad n \to \infty.$$

Suppose that $\lim_{n \to \infty} A_n A_n' = A^*$. *Then*

$$A_n X_n \xrightarrow{\mathscr{L}} N(0, A^*) \qquad \text{as} \quad n \to \infty.$$

The proofs of these lemmas are essentially the same as Lemmas 2.3, 2.4, and 2.12, respectively, in Fakhre-Zakeri [5] and will be omitted for brevity.

Now, we return to the proofs of Theorems 3.1 and 3.2. Throughout, it will be assumed that $0 < \eta < 1$ is fixed. We shall prove Theorem 3.1 by showing that $\alpha(c, \xi)$ and $\beta(c, \xi)$ converge to zero as $c \to \infty$, unif ξ.

Proof of Theorem 3.1. A lower bound for $1 - \alpha(c, \xi)$ is obtained as follows:

$$\begin{aligned} 1 - \alpha(c, \xi) &= P_\xi[\gamma \in S(\gamma_N)] \\ &\geqslant P_\xi[\|\gamma_N - \gamma\| \leqslant r(\gamma_N)] \\ &= P_\xi[U'(\gamma_N - \gamma)(\gamma_N - \gamma)'\, U \leqslant r^2(\gamma_N), \forall U\colon U'U = 1] \end{aligned} \tag{3.3}$$

Let $T_N = \operatorname{Sup}_{U'U=1} (NU'(\gamma_N - \gamma)(\gamma_N - \gamma)'\, U/U'\Sigma(\xi)\, U)$. Because of (3.3) we have

$$\begin{aligned} 1 - \alpha(c, \xi) &\geqslant P_\xi\left[T_N \leqslant N \frac{r^2(\gamma_N)}{\lambda_1(\xi)}\right] \\ &\geqslant P_\xi\left[T_N \leqslant \frac{N}{c^2\tau^2(\xi)} \cdot c^2 \cdot \frac{h^2(\xi)}{\lambda_1(\xi)} \cdot \frac{r^2(\gamma_N)}{r^2(\gamma)}\right]. \end{aligned} \tag{3.4}$$

Now, it follows from Lemmas 3.1, 3.2, and A(viii) that

$$\alpha(c, \xi) \to 0 \qquad \text{as} \quad c \to \infty, \text{unif } \xi.$$

Using (2.5) and A(v), we have

$$\begin{aligned}1-\beta(c,\xi) &= P_\xi[S(\gamma_N)\cap F(\gamma)=\varnothing]\\ &= P_\xi[\gamma_N\in B(\gamma)]\\ &\geqslant P_\xi[\gamma_N\in V(\gamma)]\\ &= P_\xi[\|\gamma_N-\gamma\|\leqslant \eta_\gamma d(\gamma)]\\ &\geqslant P_\xi[\|\gamma_N-\gamma\|\leqslant \eta_0 d(\gamma)].\end{aligned}$$

Now, by the same method as in (3.3) and (3.4) and applying Lemmas 3.1, 3.2, and A(viii), it follows that

$$\beta(c,\xi)\to 0 \qquad \text{as} \quad c\to\infty, \text{unif } \xi. \quad \blacksquare$$

Proof of Theorem 3.2. (i)

$$\begin{aligned}1-\alpha(c,\xi) &= P_\xi[\gamma\in S(\gamma_N)] = P_\xi[\gamma_N\in A(\gamma)]\\ &= P_\xi[\sqrt{N}\,(\lambda_1(\xi))^{-1/2}\,\Sigma^{1/2}(\xi)\,\Sigma^{-1/2}(\xi)(\gamma_N-\gamma)\\ &\qquad \in \sqrt{N}\,(\lambda_1(\xi))^{-1/2}\,(A(\gamma)-\gamma)].\end{aligned}$$

Write

$$\begin{aligned}\sqrt{N}\,(\lambda_1(\xi))^{-1/2}\,(A(\gamma)-\gamma) &= c\cdot\frac{\sqrt{N}}{c\tau(\xi)}\cdot(h(\xi)/\sqrt{\lambda_1(\xi)})\\ &\qquad\times(g(\xi))^{-1}\,(A(\gamma)-\gamma),\end{aligned}$$

and observe that the entries of $(\lambda_1(\xi))^{-1/2}\,\Sigma^{1/2}(\xi)$ are bounded. Now, the assertion follows from Lemmas 3.1, 3.2, and 3.3 and condition (3.1).

(ii) By the same argument as in (i), (ii) follows immediately. $\blacksquare$

4. Applications

It is the intention of this section to consider the verification of the Assumption A in the specific examples: (I) multivariate normal with unknown mean vector and unknown covariance matrix; (II) multiple regression model. We shall focus our attention on the verification of A(ii) and A(iii) where the work lies.

I (multivariate normal). Let $X_1, X_2, \ldots,$ be a sequence of multivariate normal random variables with unknown mean vector $\mu: k\times 1$ and dispersion matrix $\Sigma=(\sigma_{ij}): k\times k$, and let

$$\boldsymbol{\beta}' = (\sigma_{11}, \sigma_{12}, \ldots, \sigma_{1k}, \sigma_{22}, \ldots, \sigma_{2k}, \ldots, \sigma_{kk})' : 1\times k(k+1)/2.$$

Suppose one stops with $n+1$ observations and estimates μ by the sample mean $\bar{X}_{n+1} = (1/(n+1)) \sum_{i=1}^{n+1} X_i$ and Σ by the sample covariance matrix $S_n = (s_{ij,n}) = (1/n)\, Q$, where

$$Q = \sum_{j=1}^{n+1} (X_j - \bar{X}_{n+1})(X_j - \bar{X}_{n+1})'.$$

Let

$$\boldsymbol{\beta}_n' = (s_{11,n}, s_{12,n}, ..., s_{1k,n}, s_{22,n}, ..., s_{2k,n}, ..., s_{kk,n}) \colon 1 \times \frac{k(k+1)}{2}.$$

Obviously, $\bar{X}_{n+1}$ and $\boldsymbol{\beta}_n$ are consistent estimators of μ and $\boldsymbol{\beta}$, respectively. It is clear that A(ii) and A(iii) hold if $\gamma_n = \bar{X}_n$, for $\sqrt{n}\, \Sigma^{-1/2}(\bar{X}_n - \mu) \sim N(0, I_k)$ and is free of μ and $\boldsymbol{\beta}$. Next we show that A(ii) and A(iii) also hold if one replaces γ_n by $\boldsymbol{\beta}_n$.

Let M denote the block diagonal matrix

$$M = \operatorname{diag}(M_1, M_2, ..., M_k) \colon \frac{k(k+1)}{2} \times k^2,$$

where $M_j = [0, I_{k-j+1}] \colon (k-j+1) \times k$, $j = 1, ..., k$. Note that M has full row rank and

$$\boldsymbol{\beta}_n = M \operatorname{Vec} S_n, \qquad \boldsymbol{\beta} = M \operatorname{Vec} \Sigma, \tag{4.1}$$

where the Vec operator is the vectorization of a matrix by columns (see, e.g., Neudecker [11]). By the multivariate central limit theorem as $n \to \infty$,

$$\sqrt{n}\,(\operatorname{Vec} S_n - \operatorname{Vec} \Sigma) \xrightarrow{\mathscr{L}} N(0, V), \tag{4.2}$$

where $V = \operatorname{cov}[\operatorname{Vec}(X_1 - \mu)(X_1 - \mu)'] \colon k^2 \times k^2$. Therefore by (4.1) and (4.2) as $n \to \infty$,

$$\sqrt{n}\,(\boldsymbol{\beta}_n - \boldsymbol{\beta}) \xrightarrow{\mathscr{L}} N(0, D), \tag{4.3}$$

where $D = MVM'$. By a result of Magnus and Neudecker [10, Corollary 4.2],

$$V = (I_{k^2} + P)(\Sigma \otimes \Sigma), \tag{4.4}$$

where P is the commutation matrix and $\otimes$ denotes the Kronecker product. It is easily seen that

$$\boldsymbol{\beta}_n = M(\Sigma^{1/2} \otimes \Sigma^{1/2}) \frac{1}{n} \sum_{j=1}^{n} \operatorname{Vec} z_j z_j'$$

and (4.5)

$$\boldsymbol{\beta} = M(\Sigma^{1/2} \otimes \Sigma^{1/2}) \operatorname{Vec} I_{k(k+1)/2},$$

where z_j's are i.i.d., $N(0, I_k)$, $j = 1, 2, ..., n$. Using (4.4), we see that

$$\sqrt{n}\, D^{-1/2}(\boldsymbol{\beta}_n - \boldsymbol{\beta}) = \sqrt{n}\, D^{-1/2} M \Sigma^{1/2} \otimes \Sigma^{1/2} \left\{ \frac{1}{n} \sum_{j=1}^{n} [\text{Vec}\, z_j z_j' - \text{Vec}\, I] \right\}. \tag{4.6}$$

Thus, $\sqrt{n}\, D^{-1/2}(\boldsymbol{\beta}_n - \boldsymbol{\beta})$ can be considered as a vector of standardized sample mean, and this implies that

$$\{\sqrt{n}\, D^{-1/2}(\boldsymbol{\beta}_n - \boldsymbol{\beta});\, n \geqslant 1\} \qquad \text{is ucip.}$$

THEOREM 4.1. (a) *As* $n \to \infty$,

$$\sqrt{n}\, D^{-1/2}(\boldsymbol{\beta}_n - \boldsymbol{\beta}) \xrightarrow{\mathscr{L}} N(0, I_{k(k+1)/2}) \text{ unif } \mu, \boldsymbol{\beta};$$

(b) $\sqrt{n}\, D^{-1/2}(\boldsymbol{\beta}_n - \boldsymbol{\beta})$, $n \geqslant 1$, *is* ucip, unif μ, $\boldsymbol{\beta}$.

The following lemmas will be used for the proof of Theorem 4.1.

LEMMA 4.1 (Fakhre-Zakeri [5, Lemma 2.8]). *Let* $U = \{u \in \mathbb{R}^k : \|u\| = 1\}$. *Suppose that as* $n \to \infty$, $Y_{n\xi} \to^{\mathscr{L}} N(0, I_k)$ *uniformly in* ξ. *If* $u(\xi) \in U$, *then as* $n \to \infty$,

$$u'(\xi)\, Y_{n\xi} \xrightarrow{\mathscr{L}} N(0, 1) \text{ unif } \xi.$$

LEMMA 4.2. (*Magnus and Neudecker* [10, *Theorem* 3.1(vii), (viii)]).

$$P \,\text{Vec}\, A = \text{Vec}\, A' \tag{4.7}$$

$$P(A \otimes B) = (B \otimes A)\, P. \tag{4.8}$$

LEMMA 4.3. *Let* $\{Y_n : n \geqslant 1\}$ *be a sequence of k-dimensional random vectors which is* ucip, unif ξ, *and let* $U(\xi) = (u_{ij}(\xi))$ *be an* $m \times k$ *matrix with entries bounded on* Ξ. *Then, the sequence* $\{u(\xi)\, Y_n : n \geqslant 1\}$ *is* ucip, unif ξ.

Proof. The proof follows easily from Lemma 4.1 of Woodroofe [16], and so is omitted. ∎

Proof of Theorem 4.1 (a). Let

$$Z_n = \sqrt{n} \left[\frac{1}{n} \sum_{j=1}^{n} (\text{Vec}\, z_j z_j' - \text{Vec}\, I) \right] \tag{4.9}$$

$$U(\boldsymbol{\beta}) = D^{-1/2} M (\Sigma^{1/2} \otimes \Sigma^{1/2}). \tag{4.10}$$

It is easy to see that as $n \to \infty$,

$$Z_n \xrightarrow{\mathscr{L}} N(0, I_{k^2} + P) \qquad \text{and is free of } \mu, \boldsymbol{\beta}. \tag{4.11}$$

Since $\sqrt{n}\, D^{-1/2}(\boldsymbol{\beta}_n - \boldsymbol{\beta}) = U(\boldsymbol{\beta})\, Z_n$, it then follows from (4.3) that as $n \to \infty$,

$$U(\boldsymbol{\beta})\, Z_n \xrightarrow{\mathscr{L}} N(0, I_{k(k+1)/2}). \tag{4.12}$$

Let

$$Y_n = M Z_n. \tag{4.13}$$

In view of (4.13), one has

$$Z_n = \bar{M} Y_n, \tag{4.14}$$

where

$$
\bar{M} = \begin{array}{cc}
\left[\begin{array}{c|c|c|c|c}
I_k & 0 & 0 & \cdots & 0 \\
\hline
010\cdots0 & 0\cdots0 & 0\cdots0 & \cdots & 0 \\
0 & I_{k-1} & 0 & & 0 \\
\hline
0010\cdots0 & 0\cdots0 & 0\cdots0 & \cdots & 0 \\
0\cdots0 & 010\cdots0 & 0\cdots0 & \cdots & 0 \\
0 & 0 & I_{k-2} & & 0 \\
\hline
 & & & & \\
 & & & & \\
\hline
0\cdots01 & 0\cdots00 & 0\cdots00 & \cdots & 0 \\
0\cdots00 & 0\cdots01 & 0\cdots00 & \cdots & 0 \\
0\cdots00 & 0\cdots00 & 0\cdots1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
0\cdots00 & 0\cdots00 & 0\cdots00 & \cdots & 0 \\
0\cdots00 & 0\cdots00 & 0\cdots00 & \cdots & 0 \\
0\cdots00 & 0\cdots00 & 0\cdots00 & \cdots & 1
\end{array}\right]
&
\begin{array}{ll}
k & \text{row} \\
1 & \text{row} \\
k-1 & \text{row} \\
2 & \text{row} \\
 & \\
k-2 & \text{row} \\
 & \\
 & \\
 & \\
 & \\
 & \\
k-1 & \text{row} \\
 & \\
 & \\
1 & \text{row}
\end{array}
\\
\begin{array}{ccccc}
k & k-1 & k-2 & & 1 \\
\text{col} & \text{col} & \text{col} & & \text{col}
\end{array}
&
\end{array}
$$

$\bar{M}: k^2 \times k(k+1)/2$. In fact, $\bar{M}$ is a generalized inverse of M. Note that $M\bar{M} = I_{k(k+1)/2}$. Now, let

$$W_n = [M(I_{k^2} + P) M']^{-1/2} Y_n. \tag{4.15}$$

It follows from (4.11) that as $n \to \infty$,

$$W_n \xrightarrow{\mathscr{L}} N(0, I_{k(k+1)/2}) \qquad \text{and is free of } \mu, \boldsymbol{\beta}. \tag{4.16}$$

In view of (4.14) and (4.15), we have $U(\boldsymbol{\beta})\, Z_n = U_1(\boldsymbol{\beta})\, W_n$, where

$$U_1(\boldsymbol{\beta}) = U(\boldsymbol{\beta})\, \bar{M}[M(I_{k^2} + P) M']^{1/2}.$$

Because of (4.12) and (4.16), $U_1(\boldsymbol{\beta})$ is orthogonal. Now, the assertion (a) follows from Lemma 4.1.

Proof of Theorem 4.1 (b). Recall that $D = MVM'$. Using (4.4) and then (4.8), we get

$$D = M(\Sigma^{1/2} \otimes \Sigma^{1/2})(I_{k^2} + P)(\Sigma^{1/2} \otimes \Sigma^{1/2})\, M'. \tag{4.17}$$

Using (4.7), we have

$$\begin{aligned} \operatorname{Vec} z_j z_j' &= (\tfrac{1}{2})(I_{k^2} + P) \operatorname{Vec} z_j z_j' \\ \operatorname{Vec} I &= (\tfrac{1}{2})(I_{k^2} + P) \operatorname{Vec} I. \end{aligned} \tag{4.18}$$

After substitution of (4.18) into (4.9), we get $Z_n = (\frac{1}{2})(I_{k^2} + P)\, Z_n$. Write

$$\sqrt{n}\, D^{-1/2}(\boldsymbol{\beta}_n - \boldsymbol{\beta}) = U_2(\boldsymbol{\beta})\, Z_n,$$

where

$$U_2(\boldsymbol{\beta}) = \left(\frac{1}{2}\right) D^{-1/2} M(\Sigma^{1/2} \otimes \Sigma^{1/2})(I_{k^2} + P) : \frac{k(k+1)}{2} \times k^2.$$

Since the sequence $\{Z_n : n \geqslant 1\}$ is ucip and is free of μ, $\boldsymbol{\beta}$, if we show that each row of the matrix $U_2(\boldsymbol{\beta})$ is bounded, then the result follows from Lemma 4.3. Since $(I_{k^2} + P)(I_{k^2} + P) = 2(I_{k^2} + P)$, it follows from (4.17) that $U_2(\boldsymbol{\beta})\, U_2'(\boldsymbol{\beta}) = (\frac{1}{2})\, I_{k(k+1)/2}$. ∎

II (multiple regression model). Consider the multiple regression model

$$y_i = \boldsymbol{\beta}' \mathbf{X}^{(i)} + \varepsilon_i, \qquad i = 1, 2, ...,$$

where $\boldsymbol{\beta}' = (\beta_1, ..., \beta_k)$ is an unknown k-vector, $\mathbf{X}^{(i)} = (x_{i1}, ..., x_{ik})'$ is a k-vector of design points, $\varepsilon_1, \varepsilon_2, ...$, are i.i.d. unobservable random errors with mean zero and finite, but unknown, variance σ^2, and y_i is the observed response corresponding to the design vector $\mathbf{X}^{(i)}$. Let $\mathbf{Y}_n = (y_1, ..., y_n)'$ and $\mathbf{X}_n = (\mathbf{X}^{(1)}, ..., \mathbf{X}^{(n)})$: $k \times n$, $k \leqslant n$.

Condition B. (i) For each n the matrix $\mathbf{X}_n$ is of full rank k and there exists $\mathbf{T}$, a $k \times k$ positive definite matrix, such that

$$\frac{1}{n}(\mathbf{X}_n \mathbf{X}_n') \to \mathbf{T} \quad \text{as} \quad n \to \infty;$$

(ii) $E\,|\varepsilon_1/\sigma|^{2p} < \infty$, for some $p \geqslant 2$, for all $\boldsymbol{\beta}$ and σ.

The least squares estimate $\hat{\boldsymbol{\beta}}_n$ of $\boldsymbol{\beta}$ based on the first n observations is $\hat{\boldsymbol{\beta}}_n = (\mathbf{X}_n \mathbf{X}_n')^{-1} \mathbf{X}_n \mathbf{Y}_n$. It is well known that $\hat{\boldsymbol{\beta}}_n$ is a consistent estimator of $\boldsymbol{\beta}$. Under these conditions, we now proceed to verify the assumptions A(ii) and A(iii). We note that A(i) and A(ii) are essential to us in order to apply Anscombe's [1] theorem, that is, to show that there exists $\Sigma = \Sigma(\boldsymbol{\beta}, \sigma)$, a $k \times k$ positive definite symmetric matrix, such that

$$\sqrt{N}\,\Sigma^{-1/2}(\hat{\boldsymbol{\beta}}_N - \boldsymbol{\beta}) \xrightarrow{\mathscr{L}} N(0, I_k) \tag{4.19}$$

as $c \to \infty$ uniformly in $\boldsymbol{\beta}$, σ, or as $\xi = (\boldsymbol{\beta}, \sigma) \to \xi^*$ for any $c > 0$. It can be easily shown that the stopping time defined in (2.1) converges to ∞ a.s. as $c \to \infty$ unif ξ or as $\xi \to \xi^*$ for any $c > 0$. To prove (4.19) we shall follow the method used in Srivastava [12].

Let $\Sigma = \Sigma(\boldsymbol{\beta}, \sigma) = \sigma^2 \mathbf{T}^{-1}$ and write

$$\sqrt{n}\,\Sigma^{-/2}(\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}) = \Sigma^{-1/2}[\sigma\sqrt{n}(\mathbf{X}_n \mathbf{X}_n')^{-1/2}]\,\frac{1}{\sigma}(\mathbf{X}_n \mathbf{X}_n')^{1/2}\,(\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}).$$

To verify A(ii), since $\Sigma^{-1/2}[\sigma\sqrt{n}(\mathbf{X}_n \mathbf{X}_n')^{-1/2}] \to I_k$ as $n \to \infty$ and is free of ξ, it suffices to show that

$$\frac{1}{\sigma}(\mathbf{X}_n \mathbf{X}_n')^{1/2}\,(\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}) \xrightarrow{\mathscr{L}} N(0, I_k) \tag{4.20}$$

as $n \to \infty$ unif ξ. The left-hand side of (4.20) can be written $(\mathbf{X}_n \mathbf{X}_n')^{-1/2}\,\mathbf{X}_n \mathbf{Z}_n$, where $\mathbf{Z}_n = (1/\sigma)(\mathbf{Y}_n - E\mathbf{Y}_n)$ and its elements are i.i.d. with mean zero and variance 1. Let $\mathbf{U}_n = (\mathbf{X}_n \mathbf{X}_n')^{-1/2}\,\mathbf{X}_n = (u_{n,ij})$ and let $t_n(a) = a'\,\mathbf{U}_n \mathbf{Z}_n$, where $a \in \mathbb{R}^k$ and $a'a = 1$. Then, $t_n(a) = \sum_{j=1}^n b_{nj} z_j$, where $b_{nj} = \sum_j^k a_i u_{n,ij}$. It is easy to see that $\operatorname{Sup}_{1 \leqslant j \leqslant n} |b_{nj}|^{2+\delta} \to 0$ as $n \to \infty$ and is free of $\boldsymbol{\beta}$, σ, where $\delta \in (0, 1]$. Now, from Condition B(ii) and the Berry–Esseen theorem (see, e.g., Chow and Teicher [3, Theorem 3, p. 299]), it follows that as $n \to \infty$,

$$t_n(a) \xrightarrow{\mathscr{L}} N(0, 1) \tag{4.21}$$

unif $\boldsymbol{\beta}, \sigma$. Now, the desired result (4.19) follows from (4.21), Lemma 3.1, and a result of Srivastava [12, Corollary B_2].

To illustrate the verification of A(viii) using the asymptotic property of the stopping rule, we consider the stopping time,

$$N = \inf\left\{n : n \geqslant r_0 + c^2 \cdot \frac{(\hat{\sigma}_n^2)^{1+\delta} \lambda_n}{g(\hat{\boldsymbol{\beta}}_n)}\right\}, \tag{4.22}$$

where $\hat{\sigma}_n^2 = n^{-1}\mathbf{Y}_n'[I_n - \mathbf{X}_n'(\mathbf{X}_n\mathbf{X}_n')^{-1}\mathbf{X}_n]\mathbf{Y}_n$, $\delta > 0$, λ_n is the largest eigenvalue of $n(\mathbf{X}_n\mathbf{X}_n')^{-1}$, and g is a positive and bounded real valued function.

LEMMA 4.4. *Let $r: \mathbb{R}^k \to \mathbb{R}^k$ be such that as $y \to 0$, $r(x+y)/r(x) \to 1$ uniformly in x and let N be a stopping time satisfying* (4.22). *Then $r(\hat{\boldsymbol{\beta}}_N)/r(\boldsymbol{\beta}) \to^{\text{a.s.}} 1$ as $c \to \infty$,* unif *$\xi = (\boldsymbol{\beta}, \sigma)$ or as $\sigma \to \infty$ for any $c > 0$.*

The proof of Lemma 4.4 uses the following lemmas:

LEMMA 4.5. *Let $z_i = (1/\sigma)\,\varepsilon_i$, $i = 1, 2, \ldots$. If $\operatorname{Sup}_{\xi=(\boldsymbol{\beta},\sigma)} |z_1|^{2p} < \infty$ for some $p \geqslant 2$, then as $n \to \infty$, $\hat{\sigma}_n^2/\sigma^2 \to^{\text{a.s.}} 1$ uniformly in ξ.*

Proof. Write

$$\frac{\hat{\sigma}_n^2}{\sigma^2} = \frac{1}{n}\mathbf{Z}_n\mathbf{Z}_n' - \frac{1}{n}\mathbf{Z}_n'\mathbf{X}_n'(\mathbf{X}_n\mathbf{X}_n')^{-1}\mathbf{X}_n\mathbf{Z}_n. \tag{4.23}$$

First we show that as $n \to \infty$,

$$\frac{1}{n}\mathbf{Z}_n\mathbf{Z}_n' \xrightarrow{\text{a.s.}} 1, \tag{4.24}$$

uniformly in ξ. To prove (4.24), we need the following lemma.

LEMMA 4.6 (Ash [2, Theorem 7.2.1]). *Let $\mathbf{W}_1, \mathbf{W}_2, \ldots,$ be a sequence of independent random variables whose distribution P_ξ depends on an unknown parameter $\xi \in \Xi$ such that $E_\xi \mathbf{W}_i = 0$ for all i, and $\operatorname{Sup}_{\xi \in \Xi}[\sum_{i=1}^\infty \operatorname{Var} \mathbf{W}_i] < \infty$. Then $\sum_{i=1}^\infty \mathbf{W}_i$ converges a.s. uniformly in ξ.*

Now, let $\{b_n : n \geqslant 1\}$ be an increasing sequence of positive real numbers such that $b_n \to \infty$ as $n \to \infty$. Then, it follows from Lemma 4.6 and the Kronecker lemma that as $n \to \infty$,

$$\frac{1}{b_n}\sum_{k=1}^n b_k \mathbf{W}_k \xrightarrow{\text{a.s.}} 0,$$

uniformly in ξ. Let $\mathbf{W}_k=(z_k^2-1)/k$ and $b_k=k$, then as $n\to\infty$,

$$\frac{1}{n}\mathbf{Z}_n'\mathbf{Z}_n-1=\left(\frac{1}{n}\sum_{k=1}^{n}z_k^2\right)-1\xrightarrow{\text{a.s.}}0,\qquad \text{unif }\xi.$$

Next consider the second term on the right-hand side of (4.23); this can be written

$$\frac{1}{n}\cdot\frac{1}{\sigma^2}(\hat{\boldsymbol\beta}_n-\boldsymbol\beta)'(\mathbf{X}_n\mathbf{X}_n')(\boldsymbol\beta_n-\boldsymbol\beta)=\frac{1}{n}\cdot\frac{1}{\sigma^2}\|\hat{\boldsymbol\beta}_n-\boldsymbol\beta\|^2\cdot\sum_{i=1}^{n}\|\mathbf{X}^{(i)}\|^2.$$

By the Condition B(i), $(1/n)\sum_{i=1}^{n}\|\mathbf{X}^{(i)}\|^2$ is finite and is free of ξ and by a result of Lai and Wei [9, Corollary 2] it follows that as $n\to\infty$,

$$\frac{1}{\sigma^2}\|\beta_n-\beta\|^2\xrightarrow{\text{a.s.}}0\qquad\text{unif }\xi.\tag{4.25}$$

This completes the proof of Lemma 4.5. ∎

LEMMA 4.7. *Suppose that the stopping time $N=N_{c,\xi}$ defined in* (2.1) *satisfies the conditions,*

$$N\geqslant c^2[f(\xi)]^{2+\delta}\cdot g(\xi_N),\tag{4.26}$$

where $\delta>0$, f and g are positive real-valued functions on Ξ, and $g(\xi_N)\to^{\text{a.e.}}$ $b>0$ as $c\to\infty$, unif ξ *or as $\xi\to\xi^*$ for any $c>0$. Then $f(\xi)\sqrt{\log\log N/N}\to^{\text{a.e.}}0$ as $c\to\infty$,* unif ξ, *or as $\xi\to\xi^*$ for any $c>0$ (cf. Fakhre-Zakeri* [4, *Lemma* 5.3.2]).

We now turn to the proof of Lemma 4.4.

Proof of Lemma 4.4. It is clear that, by Lemma 4.5 and Condition B, the stopping time defined in (4.22) satisfies the condition (4.26). The assertion then follows from (4.25), a result of Lai and Wei [8, Theorem 1], and Lemma 4.7. ∎

ACKNOWLEDGMENT

The author thanks the referee for careful reading of the manuscript.

REFERENCES

[1] ANSCOMBE, F. J. (1952). Large-sample theory of sequential estimation. *Proc. Cambridge Philos. Soc.* **48** 600–607.

[2] ASH, R. B. (1972). *Real Analysis and Probability.* Academic Press, New York.

[3] CHOW, Y. S., AND TEICHER, H. (1978). *Probability Theory: Independence, Interchangeability, Martingales.* Springer-Verlag, New York.

[4] FAKHRE-ZAKERI, I. (1987). *Sequential Confidence Sets with Guaranteed Coverage Probability and Beta-Protection in Multiparameter Families.* Ph. D. thesis, University of Illinois at Urbana-Champaign.

[5] FAKHRE-ZAKERI, I. (1989). Sequential confidence sets with beta-protection in multiparameter families. *Sequential Anal.* **8** 1–26.

[6] JUHLIN, K. D. (1985). *Sequential and Non-sequential Confidence Intervals with Guaranteed Coverage Probability and Beta Protection.* Ph.D. dissertation, University of Illinois.

[7] KIM, S. L. (1987). *Sequential Confidence Sets with β-Protection in the Presence of Nuisance Parameters.* Ph.D. Dissertation, University of Illinois.

[8] LAI, T. L., AND WEI, C. Z. (1982). A law of the iterated logarithm for double arrays of independent random variables with applications to regression and time series models. *Ann. Probab.* **10** 320–335.

[9] LAI, T. L., AND WEI, C. Z. (1984). Moment inequalities with applications to regression and time series models. *Inequalities in Statistics and Probability.* IMS Lecture Notes-Monograph Series, Vol. 5, pp. 165–172. Inst. Math. Statist., Hayward, CA.

[10] MAGNUS, J. R., AND NEUDECKER, H. (1979). The commutation matrix: Some properties and applications. *Ann. Statist.* **7** 381–394.

[11] NEUDECKER, H. (1968). The Kronecker matrix product and some of its applications in econometrics. *Statist. Neerlandica* **22** 69–82.

[12] SRIVASTAVA, M. S. (1967). On fixed-width confidence bounds for regression parameters and mean vector. *J. Roy. Statist. Soc. Ser.* B 29 132–140.

[13] WIJSMAN, R. A. (1981). Confidence sets based on sequential tests. *Common. Statist.-Theor. Method A* **10** 2137–2147.

[14] WIJSMAN, R. A. (1982). Sequential confidence sets: Estimation-oriented versus test-oriented construction. *Statistical Decision Theory and Related Topics III* **2** 435–450.

[15] WIJSMAN, R. A. (1983). Sequential estimation- and test-oriented confidence intervals for μ/σ in a normal population. *Sequential Anal.* **2** 149–174.

[16] WOODROOFE, M. (1982). *Nonlinear Renewal Theory in Sequential Analysis.* CBMS-NSF Reginal Conference Series in Applied Mathematics. SIAM, Philadelphia.